

# THE BEREZIN AND GÄRDING INEQUALITIES

Y. SAFAROV

**ABSTRACT.** Let  $\varphi$  be a convex function on  $\mathbb{C}$ ,  $\mathcal{L}(\sigma)$  be a pseudodifferential operator with symbol  $\sigma$ ,  $\Lambda_\sigma$  be the set of its eigenvalues and  $m(\lambda)$  be the multiplicity of an eigenvalue  $\lambda \in \Lambda_\sigma$ . Under certain natural assumptions about properties of pseudodifferential operators, we prove that  $\sum_{\lambda \in \Lambda_\sigma} m(\lambda) \varphi(\lambda) \leq \operatorname{Re} \operatorname{Tr} \mathcal{L}(\varphi(\sigma)) + R$ , where  $R$  is an error term of the same order as the remainder term in the Gårding inequality.

## 1. INTRODUCTION

Let  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  be a convex function,  $B$  be a self-adjoint operator and  $P$  be an orthogonal projection in a separable Hilbert space  $H$ . Then

$$(1) \quad \operatorname{Tr} \varphi(P B|_{PH}) \leq \operatorname{Tr} (P \varphi(B)|_{PH}),$$

provided that the operators in the right and left hand sides are well defined and belong to the trace class  $\mathfrak{S}_1$ . This estimate was proved in [1] and is often called “the Berezin inequality” (some generalizations of (1) were obtained in [8]). By the spectral theorem, every self-adjoint operator is unitary equivalent to a multiplication operator in  $L_2$ . Therefore Berezin’s result can be reformulated in the following way: if  $\varphi : \mathbb{R} \rightarrow \mathbb{R}$  is a convex function,  $b$  is a real-valued function,  $\{b\}$  is the corresponding multiplication operator,  $U : H \rightarrow L_2$  is an isometry onto a subspace of  $L_2$  and  $\mathcal{Q}(b) := U^* \{b\} U$ , then

$$(2) \quad \operatorname{Tr} \varphi(\mathcal{Q}(b)) \leq \operatorname{Tr} \mathcal{Q}(\varphi(b))$$

whenever the operators  $\mathcal{Q}(\sigma)$  and  $\mathcal{Q}(\varphi(b))$  are well defined and belong to  $\mathfrak{S}_1$ .

The Berezin inequality has been used for the study of spectral properties of differential and pseudodifferential operators. If  $B$  is a self-adjoint pseudodifferential operator with symbol  $\sigma_B$  then, under certain assumptions,  $\varphi(B)$  is a pseudodifferential operator whose symbol coincides with  $\varphi(\sigma_B)$  modulo a lower order term. In this case the right hand side of (1) is equal to the sum of an integral of  $\varphi(\sigma_B)$  and a lower order remainder, and (1) implies asymptotic formulas for the spectrum of the operator  $P B|_{PH}$  (see, for example, [9]). If there exist an isometry  $U : H \rightarrow L_2$  and a function  $b$  such that  $B = \mathcal{Q}(b)$ , and  $\operatorname{Tr} \mathcal{Q}(\varphi(b))$  is given by an explicit formula then (2) yields estimates for the spectrum of the operator  $B$  itself (see, for instance, [4]).

---

*Date:* December 2004.

*1991 Mathematics Subject Classification.* 47A63, 47G30.

*Key words and phrases.* Convex functions, operator inequalities.

The first scheme works only for self-adjoint pseudodifferential operators  $B$  and relies on symbolic functional calculus. The second allows one to obtain estimates only in terms of the function  $b$  which depends on the choice of the isometry  $U$ . The main problem in this scheme is to construct a suitable isometry  $U$  and to investigate the relation between  $b$  and the actual symbol  $\sigma_B$ . For some operators this can be done with the use of the so-called coherent states (as in [2] or [4]). With the exception of some very special cases, the formulas relating  $b$  and  $\sigma_B$  contain a lower order error term. These formulas, together with (2), imply asymptotic estimates in terms of  $\sigma_B$  with a similar error term.

The aim of this paper is to show that the Berezin inequality is an elementary consequence of the Gårding inequality. If the Gårding inequality holds with a lower order error term then the Berezin inequality contains an error term of the same order. Note that the Gårding inequality is a simpler result than a symbolic functional calculus or a coherent state representation and, as a rule, immediately follows from either of these two.

The inequalities (1) and (2) are easily proved by representing the quadratic form of  $P B|_{PH}$  or  $B$  as a Lebesgue integral and applying Jensen's inequality. Our proof does not involve the spectral theorem or Lebesgue integrals. Instead, we observe that Jensen's inequality holds for more general functionals and apply it to suitably chosen functionals on the space of symbols.

## 2. CONVEX FUNCTIONS

In this section we shall briefly recall some results from convex analysis. Keeping in mind possible applications to operator-valued functions  $\sigma$ , we shall consider convex functions on an infinite dimensional locally convex real vector space  $X$ . All results and their proofs are elementary and cannot be substantially simplified even if  $X = \mathbb{R}$ .

A function  $\varphi : X \rightarrow [-\infty, +\infty]$  is called *convex* if its epigraph  $\mathcal{E}(\varphi) := \{(t, \mathbf{x}) \in \mathbb{R} \times X : t \geq \varphi(\mathbf{x})\}$  is a convex subset of  $\mathbb{R} \times X$ . A convex function is said to be *proper* if  $|\varphi| \not\equiv \infty$  and *closed* (or lower semicontinuous) if  $\mathcal{E}(\varphi)$  is closed in  $\mathbb{R} \times X$  in the product topology. If  $\mathcal{E}(\varphi)$  lies to one side of a hyperplane  $\mathcal{H}$  passing through the point  $(\varphi(\mathbf{x}), \mathbf{x})$  then  $\mathcal{H}$  is said to be a *supporting* hyperplane at  $(\varphi(\mathbf{x}), \mathbf{x})$ . If  $\varphi$  is convex and  $\dim X < \infty$  then each point  $(\varphi(\mathbf{x}), \mathbf{x})$  has at least one (possibly, vertical) supporting hyperplane. In the infinite dimensional case there may be no supporting hyperplanes.

**Example 1.** Let  $X = \mathbb{R}^\infty$  be the space of real sequences  $\mathbf{x} = \{x_1, x_2, \dots\}$  provided with the topology of element-wise convergence,  $\varphi(\mathbf{x}) := \sum_{j=1}^\infty x_j^{-1}$  if  $x_j > 0$  for all  $j$  and the sum is finite, and  $\varphi(\mathbf{x}) := +\infty$  otherwise. Then the closed convex set  $\mathcal{E}(\varphi)$  does not have any supporting hyperplanes; in other words, no linear continuous functional attains its minimal value on  $\mathcal{E}(\varphi)$ .

If the supporting hyperplane  $\mathcal{H}$  is not vertical then it coincides with the graph of an affine function  $l_{t^*, \mathbf{x}^*}(\mathbf{x}) := \langle \mathbf{x}^*, \mathbf{x} \rangle - t^*$  with some  $t^* \in \mathbb{R}$  and  $\mathbf{x}^*$

from the dual space  $X^*$ . The set of vectors  $\mathbf{x}^* \in X^*$  generating non-vertical supporting hyperplanes at  $(\varphi(\mathbf{x}), \mathbf{x})$  is called the *subdifferential* of  $\varphi$  at the point  $\mathbf{x}$  and is denoted by  $\partial\varphi(\mathbf{x})$ .

The closed convex function  $\varphi^*(\mathbf{x}^*) := \sup_{\mathbf{x} \in X} \{\langle \mathbf{x}^*, \mathbf{x} \rangle - \varphi(\mathbf{x})\}$  on  $X^*$  is called the *conjugate* of  $\varphi$ . The following well known result (see, for example, [5] or [6]) is a simple consequence of the separation theorem.

**Lemma 2.** *If  $\varphi$  is a proper closed convex function then  $\varphi \equiv \varphi^{**}$ .*

We have  $(t^*, \mathbf{x}^*) \in \mathcal{E}(\varphi^*)$  if and only if  $l_{t^*, \mathbf{x}^*}(\mathbf{x}) \leq \varphi(\mathbf{x})$  for all  $\mathbf{x} \in X$ . Given  $\varepsilon \geq 0$  and  $\mathbf{x}_0 \in X$ , let us denote by  $\partial_\varepsilon \varphi(\mathbf{x}_0)$  the set of points  $(t^*, \mathbf{x}^*) \in \mathcal{E}(\varphi^*)$  such that  $\varphi(\mathbf{x}_0) - \varepsilon \leq l_{t^*, \mathbf{x}^*}(\mathbf{x}_0)$ . In particular,

$$\partial_0 \varphi(\mathbf{x}_0) = \{(t^*, \mathbf{x}^*) \in \mathcal{E}(\varphi^*) : \mathbf{x}^* \in \partial\varphi(\mathbf{x}_0), t^* = \langle \mathbf{x}^*, \mathbf{x}_0 \rangle - \varphi(\mathbf{x}_0)\}.$$

The set  $\partial_0 \varphi(\mathbf{x}_0)$  may well be empty even in the case  $X = \mathbb{R}$ ; if  $\dim X = \infty$  then it may happen that  $\partial_0 \varphi(\mathbf{x}_0) = \emptyset$  for all  $\mathbf{x}_0 \in X$  (see Example 1). However, by Lemma 2,  $\partial_\varepsilon \varphi(\mathbf{x}_0) \neq \emptyset$  for each  $\varepsilon > 0$  provided that  $\varphi(\mathbf{x}_0) < +\infty$ .

### 3. JENSEN'S INEQUALITY

Let  $L_X$  be a set of functions  $\sigma : \Xi \rightarrow X$  defined on a nonempty set  $\Xi$ ,  $L_{\mathbb{R}}$  be a set of real-valued functions on  $\Xi$ ,  $\mathfrak{I}_X$  be a map from  $L_X$  to  $X$  and  $\mathfrak{I}_{\mathbb{R}} : L_{\mathbb{R}} \rightarrow [-\infty, +\infty]$  be a real functional on  $L_{\mathbb{R}}$ .

**Lemma 3.** *Let  $\varphi$  be a proper closed convex function such that  $\varphi(\sigma) \in L_{\mathbb{R}}$ . Assume that for each  $\varepsilon > 0$  there exists a point  $(t^*, \mathbf{x}^*) \in \partial_\varepsilon \varphi(\mathfrak{I}_X(\sigma))$  such that*

- (a<sub>1</sub>)  $l_{t^*, \mathbf{x}^*}(\sigma) \in L_{\mathbb{R}}$  and  $l_{t^*, \mathbf{x}^*}(\mathfrak{I}_X(\sigma)) \leq \mathfrak{I}_{\mathbb{R}}(l_{t^*, \mathbf{x}^*}(\sigma)) + C_1$ ,
- (a<sub>2</sub>)  $\mathfrak{I}_{\mathbb{R}}(l_{t^*, \mathbf{x}^*}(\sigma)) \leq \mathfrak{I}_{\mathbb{R}}(\varphi(\sigma)) + C_2$ ,

where  $C_1$  and  $C_2$  are real constants. Then  $\varphi(\mathfrak{I}_X(\sigma)) \leq \mathfrak{I}_{\mathbb{R}}(\varphi(\sigma)) + C_1 + C_2$ .

*Proof.* Since  $(t^*, \mathbf{x}^*) \in \partial_\varepsilon \varphi(\mathfrak{I}_X(\sigma))$ , the conditions (a<sub>1</sub>) and (a<sub>2</sub>) imply that  $\varphi(\mathfrak{I}_X(\sigma)) - \varepsilon \leq l_{t^*, \mathbf{x}^*}(\mathfrak{I}_X(\sigma)) \leq \mathfrak{I}_{\mathbb{R}}(l_{t^*, \mathbf{x}^*}(\sigma)) + C_1 \leq \mathfrak{I}_{\mathbb{R}}(\varphi(\sigma)) + C_1 + C_2$ . Letting  $\varepsilon \rightarrow 0$ , we obtain the required inequality.  $\square$

The inequality (a<sub>1</sub>) holds with  $C_1 = 0$  provided that

- (a'<sub>1</sub>) the functional  $\mathfrak{I}_{\mathbb{R}}$  is linear,  $\mathfrak{I}_{\mathbb{R}}(1) = 1$  and  $\mathfrak{I}_{\mathbb{R}}(\langle \mathbf{x}^*, \sigma \rangle) = \langle \mathbf{x}^*, \mathfrak{I}_X(\sigma) \rangle$ .

The condition (a<sub>2</sub>) is fulfilled with  $C_2 = 0$  for all monotone functionals  $\mathfrak{I}_{\mathbb{R}}$ , that is, the functionals  $\mathfrak{I}_{\mathbb{R}}$  satisfying

- (a'<sub>2</sub>)  $\mathfrak{I}_{\mathbb{R}}(\sigma_1) \leq \mathfrak{I}_{\mathbb{R}}(\sigma_2)$  whenever  $\sigma_1, \sigma_2 \in L_{\mathbb{R}}$  and  $\sigma_1(\theta) \leq \sigma_2(\theta)$  for all  $\theta \in \Xi$ .

If  $\mathfrak{I}_{\mathbb{R}}$  is the Lebesgue integral with respect to a probability measure or the normalized Perron integral and  $\mathfrak{I}_X$  is the corresponding vector-valued integral understood in the weak sense then (a'<sub>1</sub>) and (a'<sub>2</sub>) hold for all integrable functions and all  $(t^*, \mathbf{x}^*) \in \mathbb{R} \times X^*$ . Therefore Lemma 3 implies the standard Jensen's inequality. The following example is less obvious.

**Example 4.** Let  $A$  be a linear positive operator in  $L_2(\Xi)$ ,  $L_{\mathbb{R}}$  be the space of measurable bounded functions on  $\Xi$  and  $V_{\sigma}$  be the operator of multiplication by the function  $\sigma$ . Denote by  $\lambda_1(\sigma), \lambda_2(\sigma), \dots$  the ordered eigenvalues of the Friedrichs extension of the operator  $A + V_{\sigma}$  lying below its essential spectrum.

Let us fix  $n \in \mathbb{N}$  and take  $X = \mathbb{R}$  and  $\mathfrak{I}_{\mathbb{R}} = \lambda_n(\sigma)$ . In view of the Rayleigh–Ritz formula,  $\mathfrak{I}_{\mathbb{R}}$  satisfies  $(\mathbf{a}'_2)$  for all  $\sigma \in L_{\mathbb{R}}$ . We have

$$A + tV_{\sigma} \geq \delta(A + V_{\sigma}) + (t - \delta)\lambda_n(\sigma) + \inf_{\theta \in \Xi} ((t - \delta)(\sigma(\theta) - \lambda_n))$$

for all  $\delta \in [0, 1]$  and  $t \in \mathbb{R}$ . Let  $\varphi$  be a proper closed convex function such that  $\varphi(\sigma) \in L_{\mathbb{R}}$ . The above inequality implies  $(\mathbf{a}_1)$  with

$$C_1 = F(\lambda_n(\sigma)) := \max \{a_-(\sup \sigma - \lambda_n(\sigma))_+, (b - 1)_+(\lambda_n(\sigma) - \inf \sigma)\},$$

where  $a = \sup \partial\varphi(\inf \sigma)$ ,  $b = \inf \partial\varphi(\sup \sigma)$ , and the subscripts  $\pm$  denote the positive and negative parts (we define  $\sup \partial\varphi(t) := -\infty$  and  $\inf \partial\varphi(t) := +\infty$  when  $\partial\varphi(t) = \emptyset$ ). Therefore, by Lemma 3,

$$\varphi(\lambda_n(\sigma)) \leq \lambda_n(\varphi(\sigma)) + F(\lambda_n(\sigma)).$$

Note that  $F(\lambda_n(\sigma)) = 0$  whenever  $[a, b] \subset [0, 1]$  or  $[a, b] \subset [-\infty, 1]$  and  $\sup \sigma \leq \lambda_n(\sigma)$ .

#### 4. BEREZIN INEQUALITY

Let  $L_{\mathbb{C}}$  be a linear space of complex-valued functions on  $\Xi$  containing the constant functions and closed with respect to the complex conjugation, and let  $L_{\mathbb{R}}$  be the subspace of real-valued functions. Consider a linear map  $\mathcal{Q}$  from  $L_{\mathbb{C}}$  into the space of linear operators in a separable Hilbert space  $H$  (a quantization) such that  $\mathcal{Q}(1) = I$  and  $\mathcal{D}(\mathcal{Q}(\sigma)) = \mathcal{D}(\mathcal{Q}(\operatorname{Re} \sigma)) \cap \mathcal{D}(\mathcal{Q}(\operatorname{Im} \sigma))$ . If  $\sigma \in L_{\mathbb{C}}$ , let  $\Omega_{\sigma}$  be the numerical range of the operator  $\mathcal{Q}(\sigma)$ ,  $\Lambda_{\sigma}$  be the set of its eigenvalues and  $m(\lambda)$  be the algebraic multiplicity of the eigenvalue  $\lambda \in \Lambda_{\sigma}$ .

Let us fix a bounded operator  $T$  and, given  $\nu \in \mathbb{R}$ , denote by  $G_{\nu}$  the set of functions  $\sigma \in L_{\mathbb{R}}$  such that

$$(\mathbf{G}) \quad \operatorname{Re}(\mathcal{Q}(\sigma)u, u)_H \geq -\nu(Tu, u)_H \text{ for all } u \in \mathcal{D}(\mathcal{Q}(\sigma)).$$

**Lemma 5.** *Let  $\sigma \in L_{\mathbb{C}}$  and  $\varphi$  be a proper closed convex function on  $\mathbb{C}$ . Assume that for each  $\varepsilon > 0$  and each  $z \in \Omega_{\sigma}$  there exists  $(t^*, z^*) \in \partial_{\varepsilon}\varphi(z)$  satisfying the following two conditions:*

- (b<sub>1</sub>)  $\operatorname{Im}(\mathcal{Q}(\operatorname{Im}(z^*\bar{\sigma}))u, u)_H \leq \nu_1(Tu, u)_H$  for all  $u \in \mathcal{D}(\mathcal{Q}(\sigma))$ ,
- (b<sub>2</sub>)  $\varphi(\sigma) - \operatorname{Re}(z^*\bar{\sigma}) + t^* \in G_{\nu_2}$ ,

where  $\nu_1$  and  $\nu_2$  are real constants. Then

$$(3) \quad \varphi((\mathcal{Q}(\sigma)u, u)_H) \leq \operatorname{Re}(\mathcal{Q}(\varphi(\sigma))u, u)_H + (\nu_1 + \nu_2)(Tu, u)_H$$

whenever  $u \in \mathcal{D}(\mathcal{Q}(\sigma)) \cap \mathcal{D}(\mathcal{Q}(\varphi(\sigma)))$  and  $\|u\|_H = 1$ .

*Proof.* Let us identify  $\mathbb{C}$  with  $\mathbb{R}^2$  so that  $\langle z^*, z \rangle = \operatorname{Re}(z^* \bar{z})$ ,  $\forall z, z^* \in \mathbb{C}$ . Then (3) immediately follows from Lemma 3 with  $\mathfrak{I}_{\mathbb{C}}(\sigma) = (\mathcal{Q}(\sigma)u, u)_H$ ,  $\mathfrak{I}_{\mathbb{R}}(\sigma) = \operatorname{Re}(\mathcal{Q}(\sigma)u, u)_H$  and  $C_j = \nu_j(Tu, u)_H$ .  $\square$

**Lemma 6.** *Let  $\sigma \in L_{\mathbb{R}}$  and  $\varphi$  be a proper closed convex function on  $\mathbb{R}$ . If*

*(b'<sub>1</sub>) the operator  $\mathcal{Q}(\sigma)$  is symmetric*

*and for each  $\varepsilon > 0$  and  $z \in \Omega_{\sigma}$  there exists  $(t^*, z^*) \in \partial_{\varepsilon}\varphi(z)$  satisfying (b<sub>2</sub>) then (3) holds with  $\nu_1 = 0$ .*

*Proof.* This is a particular case of Lemma 5 with the convex function on  $\mathbb{C}$  which is equal to  $\varphi$  on  $\mathbb{R}$  and to  $+\infty$  on  $\mathbb{C} \setminus \mathbb{R}$ .  $\square$

Note that the second condition in Lemmas 5 and 6 is satisfied whenever  $\varphi$  is differentiable and

*(b'<sub>2</sub>)  $\varphi(\sigma) - \varphi(z) - \operatorname{Re}(\varphi'_z(z)(\sigma - z)) \in G_{\nu_2}$  for all  $z \in \Omega_{\sigma}$ .*

**Theorem 7.** *Assume that  $\sigma$  and  $\varphi$  satisfy the conditions of Lemma 5 or Lemma 6,  $\mathcal{Q}(\varphi(\sigma)) \in \mathfrak{S}_1$ ,  $T \in \mathfrak{S}_1$ , and at least one of the following two conditions is fulfilled:*

*(c<sub>1</sub>)  $\varphi$  is nonnegative;*

*(c<sub>2</sub>) the set of the generalized eigenvectors of the operator  $\mathcal{Q}(\sigma)$  is complete.*

*Then the set  $\Lambda_{\sigma, \varphi}^+ := \{\lambda \in \Lambda_{\sigma} : \varphi(\lambda) > 0\}$  is countable, each eigenvalue  $\lambda \in \Lambda_{\sigma, \varphi}^+$  has a finite algebraic multiplicity,  $\sum_{\lambda \in \Lambda_{\sigma, \varphi}^+} m(\lambda) \varphi(\lambda) < \infty$  and*

$$(4) \quad \sum_{\lambda \in \Lambda_{\sigma}} m(\lambda) \varphi(\lambda) \leq \operatorname{Re} \operatorname{Tr} \mathcal{Q}(\varphi(\sigma)) + (\nu_1 + \nu_2) \operatorname{Tr} T.$$

*Proof.* Let  $\mathbf{v} = \{v_1, v_2, \dots, v_k\}$  be a finite collection of the generalized eigenvectors of the operator  $\mathcal{Q}(\sigma)$  corresponding to the eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_k$ . Denote by  $H_{\mathbf{v}}$  the finite dimensional invariant subspace of  $\mathcal{Q}(\sigma)$  spanned by the vectors  $(\mathcal{Q}(\sigma) - \lambda_j)^n v_j$ ,  $j = 1, \dots, k$ ,  $n = 0, 1, \dots$ . The restriction  $\mathcal{Q}(\sigma)|_{H_{\mathbf{v}}}$  has the same eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_k$  whose algebraic multiplicities  $m_{\mathbf{v}}(\lambda_j)$  do not exceed  $m(\lambda_j)$ . By Schur's lemma, the operator  $\mathcal{Q}(\sigma)|_{H_{\mathbf{v}}}$  is represented by a triangular matrix with the diagonal entries  $\lambda_j$  in some orthonormal basis  $\{u_j\} \subset H_{\mathbf{v}}$ . Applying (3) to the vectors  $u_j$ , we see that

$$(5) \quad \sum_{j=1}^k m_{\mathbf{v}}(\lambda_j) \varphi(\lambda_j) \leq \operatorname{Re} \operatorname{Tr} \Pi_{\mathbf{v}} \mathcal{Q}(\varphi(\sigma)) + (\nu_1 + \nu_2) \operatorname{Tr} \Pi_{\mathbf{v}} T$$

$$\leq \|\mathcal{Q}(\varphi(\sigma))\|_{\mathfrak{S}_1} + \|(\nu_1 + \nu_2) T\|_{\mathfrak{S}_1},$$

where  $\Pi_{\mathbf{v}}$  is the orthogonal projection onto the subspace  $H_{\mathbf{v}}$ .

If the set  $\Lambda_{\sigma, \varphi}^+$  were uncountable or there were an eigenvalue  $\lambda \in \Lambda_{\sigma, \varphi}^+$  of infinite algebraic multiplicity then we could find a positive constant  $\delta$  and an arbitrarily large collection  $\mathbf{v}$  of eigenvectors  $v_j$  such that  $\lambda_j \in \Lambda_{\sigma, \varphi}^+$  and  $\varphi(\lambda_j) \geq \delta$ . This contradicts to (5). Therefore  $\Lambda_{\sigma, \varphi}^+$  is countable and  $m(\lambda) < \infty$  for all  $\lambda \in \Lambda_{\sigma, \varphi}^+$ . Choosing a sequence of expanding finite sets  $\mathbf{v}_1 \subset \mathbf{v}_2 \subset$

$\mathbf{v}_3 \subset \dots$  whose union  $\bigcup_{n=1}^{\infty} \mathbf{v}_n$  contains all eigenvectors corresponding to the eigenvalues  $\lambda \in \Lambda_{\sigma, \varphi}^+$ , applying (5) with  $\mathbf{v} = \mathbf{v}_n$  and letting  $n \rightarrow \infty$ , we see that  $\sum_{\lambda \in \Lambda_{\sigma, \varphi}^+} m(\lambda) \varphi(\lambda)$  is estimated by the same sum of the trace norms.

In order to prove (4), we note that the left hand side of (4) is  $-\infty$  whenever the set  $\Lambda_{\sigma, \varphi}^- := \{\lambda \in \Lambda_{\sigma} : \varphi(\lambda) < 0\}$  is uncountable or contains an eigenvalue of infinite algebraic multiplicity. Therefore we can assume without loss of generality that the set of generalized eigenvectors corresponding to the eigenvalues  $\lambda \in \Lambda_{\sigma, \varphi}^+ \cup \Lambda_{\sigma, \varphi}^-$  is countable. Let us choose a sequence of finite sets  $\mathbf{v}_1 \subset \mathbf{v}_2 \subset \mathbf{v}_3 \subset \dots$  such that  $\bigcup_{n=1}^{\infty} \mathbf{v}_n$  contains all these eigenvectors and  $\overline{\bigcup_{n=1}^{\infty} H_{\mathbf{v}_n}} = H'$ , where  $H'$  denotes the closed linear span of all generalized eigenvectors of the operator  $\mathcal{Q}(\sigma)$ .

If  $(\mathbf{c}_2)$  holds then  $H' = H$ . Therefore, taking  $\mathbf{v} = \mathbf{v}_n$  in the first inequality (5) and letting  $n \rightarrow \infty$ , we arrive at (4). If  $H' \neq H$  and  $\varphi \geq 0$  then we choose an orthonormal basis  $\{\tilde{u}_i\}$  in  $(H')^{\perp}$ , apply (3) to  $\tilde{u}_i$  and add up the obtained inequalities and the first inequality (5) with  $\mathbf{v} = \mathbf{v}_n$ . Since  $\varphi((\mathcal{Q}(\sigma)\tilde{u}_i, \tilde{u}_i)_H) \geq 0$ , now (4) is proved by letting  $n \rightarrow \infty$ .  $\square$

If a convex function  $\varphi : \mathbb{R} \rightarrow [-\infty, +\infty]$  takes negative values then the set of its zeros consists of at most two points. In this case the inclusion  $\varphi(\mathcal{Q}(\sigma)) \in \mathfrak{S}_1$  implies  $(\mathbf{c}_2)$  for each self-adjoint operator  $\mathcal{Q}(\sigma)$ . Therefore (2) is a particular case of Theorem 7 with  $T = 0$ .

## 5. PSEUDODIFFERENTIAL OPERATORS

In the theory of pseudodifferential operators,  $\Xi$  is the cotangent bundle  $T^*M$  over a domain  $M \subset \mathbb{R}^n$  or a manifold  $M$  and  $H = L_2(M)$ . For  $M \subset \mathbb{R}^n$ , quantization is defined by the formula

$$(6) \quad \mathcal{Q}_{\tau}(\sigma)u(x) = (2\pi h)^{-n} \int_{\mathbb{R}^n} \int_M e^{ih^{-1}(x-y) \cdot \xi} \sigma(\tau x + (1-\tau)y, \xi) u(y) dy d\xi,$$

where  $\tau$  is a fixed number from the interval  $[0, 1]$  and  $h$  is a real parameter (see, for example, [3]). This definition can be extended to manifolds  $M$  (see [10]). The functions  $\sigma$  on  $T^*M$  are called  $\tau$ -symbols; for  $\tau = \frac{1}{2}$  they called Weyl symbols. In the classical theory of pseudodifferential operators one takes  $h = 1$  and defines the order of a symbol  $\sigma$  in terms of its behaviour for large  $\xi$ . In the semiclassical theory the order is defined in terms of asymptotic behaviour as  $h \rightarrow 0$ .

If  $M \subset \mathbb{R}^n$  then, obviously,  $(\mathcal{Q}_{\tau}(\sigma)u, u)_H = (u, \mathcal{Q}_{1-\tau}(\bar{\sigma})u)_H$ . This equality remains true on a manifold  $M$  if the  $\tau$ -quantization is defined as in [10]. It implies that the estimate  $(\mathbf{b}_1)$  holds for  $\mathcal{Q} = \mathcal{Q}_{1/2}$  with  $\nu_1 = 0$ .

If  $\sigma \geq 0$  and  $T$  is a lower order operator then  $(\mathbf{G})$  is known as the Gårding inequality or the Fefferman–Phong inequality (the latter gives a more precise result in terms of the order of  $T$ ). The constant  $\nu$  in this inequality can usually be estimated by a functional  $\mathcal{G}(\sigma)$  which involves partial derivatives of the symbol  $\sigma$  up to a certain order (see, for example, [3], [7], [12] or [13]).

The condition  $(\mathbf{b}_2)$  means that  $(\mathbf{G})$  holds uniformly on the set of nonnegative functions  $\varphi(\sigma) - \operatorname{Re}(z^* \bar{\sigma}) + t^*$ , which is the case whenever the functional  $\mathcal{G}$  is uniformly bounded on this set. One can obtain explicit formulas for  $\mathcal{G}$  by analyzing the known proofs of the Gårding inequality. However, such analysis lies outside the scope of this paper. Instead, we conclude by giving three examples which demonstrate possible applications of Theorem 7 (in the last two examples  $(\mathbf{b}_2)$  can be proved directly).

**Example 8.** Let  $M$  be a compact  $n$ -dimensional  $C^\infty$ -manifold,  $H$  be the space of square integrable half-densities on  $M$ ,  $S^m$  be the Hörmander class of symbols and  $\Psi^m$  be the corresponding class of classical pseudodifferential operators in  $H$ . Consider an elliptic positive pseudodifferential operator  $A \in \Psi^1$  and denote by  $\Pi_\mu$  its spectral projection corresponding to the interval  $[0, \mu)$ . It is well known that  $\operatorname{rank} \Pi_\mu < \infty$  and

$$(7) \quad \operatorname{Tr}(B \Pi_\mu) = (2\pi)^{-n} \int_{\sigma_A(x, \xi) < \mu} \sigma_B(x, \xi) \, dx \, d\xi + O(\mu^{n+m-1}), \quad \mu \rightarrow \infty,$$

for every  $B \in \Psi^m$  provided that  $n+m-1 \geq 0$ , where  $\sigma_A$  and  $\sigma_B$  are principal symbols of the operators  $A$  and  $B$  (see, for example, [7] or [11]).

Let  $L_{\mathbb{C}} = S^0$ , the quantization  $\mathcal{Q}_{1/2} : S^0 \rightarrow \Psi^0$  be defined as in [10],  $\mathcal{Q}(\sigma) := \Pi_\mu \mathcal{Q}_{1/2}(\sigma)|_{\Pi_\mu H}$  and  $T = \Pi_\mu A^{-1}|_{\Pi_\mu H}$ . If  $\varphi \in C^\infty(\mathbb{C})$  is a convex function then, by the Gårding inequality, we have  $(\mathbf{b}'_2)$  with some constant  $\nu_2$  depending on  $\sigma$  and  $\varphi$ . Since  $\sigma - \sigma_B \in S^{-1}$  whenever  $B = \mathcal{Q}_{1/2}(\sigma)$  (see [10]), Theorem 7 and (7) imply that

$$\sum_{\lambda \in \Lambda_{\mu, B}} m(\lambda) \varphi(\lambda) \leq (2\pi)^{-n} \int_{\sigma_A(x, \xi) < \mu} \varphi(\sigma_B(x, \xi)) \, dx \, d\xi + O(\mu^{n-1}), \quad \mu \rightarrow \infty$$

for every operator  $B \in \Psi^0$ , where  $\Lambda_{\mu, B}$  is the set of eigenvalues of  $\Pi_\mu B|_{\Pi_\mu H}$  and  $m(\lambda)$  is the algebraic multiplicity of  $\lambda$ .

*Remark 9.* The above inequality was obtained in [9] for self-adjoint operators  $B \in \Psi^0$ .

In the following examples  $M \subset \mathbb{R}^n$  is an open bounded set and  $D$  is an open bounded subset of  $T^*M$ .

**Example 10.** Let  $\sigma$  be the characteristic function of  $D$ ,  $\sigma' := 1 - \sigma$ ,  $L_{\mathbb{C}}$  be the linear space spanned by  $\sigma$  and  $\sigma'$ , and let  $\mathcal{Q} = \mathcal{Q}_1$ . If  $R := \mathcal{Q}_1(\sigma) \mathcal{Q}_0(\sigma')$  then  $\mathcal{Q}_1(\sigma) = \mathcal{Q}_1(\sigma) \mathcal{Q}_0(\sigma) + R$ ,  $\mathcal{Q}_1(\sigma') = \mathcal{Q}_1(\sigma') \mathcal{Q}_0(\sigma') + R^*$  and

$$2 \operatorname{Im}(\mathcal{Q}_1(\operatorname{Im}(z^* \sigma))u, u)_H = \operatorname{Im} z^*((R - R^*)u, u)_H, \quad \forall z^* \in \mathbb{C}.$$

Therefore the conditions  $(\mathbf{b}_1)$  and  $(\mathbf{b}'_2)$  are satisfied with  $T = |\operatorname{Re} R| + |\operatorname{Im} R|$ ,  $\nu_1 = \sup_{z \in \Omega_\sigma} |\operatorname{Im} \varphi'_z(z)|$  and  $\nu_2 = \sup_{z \in \Omega_\sigma} |\varphi(1) - \operatorname{Re} \varphi'_z(z)|$  for every nonnegative differentiable convex function  $\varphi : \mathbb{C} \rightarrow \mathbb{R}$  vanishing at the origin  $z = 0$ .

The operators  $\mathcal{Q}_1(\sigma)$  and  $\mathcal{Q}_0(\sigma')$  belong to the Hilbert-Schmidt class  $\mathfrak{S}_2$ . Therefore  $T \in \mathfrak{S}_1$ ,  $\mathcal{Q}_1(\varphi(\sigma)) \in \mathfrak{S}_1$ , and Theorem 7 implies that

$$\sum_{\lambda \in \Lambda_\sigma} m(\lambda) \varphi(\lambda) \leq (2\pi h)^{-n} \varphi(1) \int_D dx d\xi + 4 \left( \varphi(1) + \sup_{z \in \Omega_\sigma} |\varphi'_z(z)| \right) \|R\|_{\mathfrak{S}_1}.$$

Approximating  $\sigma$  and  $\sigma'$  by smooth functions, which are supported in  $D$  and  $T^*M \setminus D$  respectively and vanish near the boundary  $\partial D$ , we see that  $\|R\|_{\mathfrak{S}_1} = o(h^{-n})$  as  $h \rightarrow 0$ . If the Minkowski dimension  $d$  of  $\partial D$  is strictly smaller than  $n$  then one can improve this estimate and show that  $\|R\|_{\mathfrak{S}_1} = O(h^{\alpha-n})$  where  $\alpha$  is a positive constant depending on  $d$ .

**Example 11.** Let  $C_0^\kappa(D)$  be the subspace of the Hölder space  $C^\kappa(\mathbb{R}^{2n})$  which consists of real-valued function vanishing outside  $D$ . Let us fix  $\varepsilon > 0$ , denote by  $L_\mathbb{R}$  the real linear space spanned by  $C_0^{n/2+2+\varepsilon}(D)$  and constant functions, and consider the Weyl quantization on  $L_\mathbb{C} = L_\mathbb{R}$ .

Let  $\sigma \in C_0^{n/2+2+\varepsilon}(D)$  and  $\varphi$  be a nonnegative convex function on the closure  $\overline{\Omega_\sigma}$  such that  $\varphi \in C^{n/2+2+\varepsilon}$  and  $\varphi'' \geq \delta > 0$  (recall that  $\Omega_\sigma \subset \mathbb{R}$ ). Then  $\varphi_z(\sigma) := \varphi(\sigma) - \varphi(z) - \varphi'(z)(\sigma - z) = (\psi_z(\sigma))^2$  where  $\psi_z \in C^{n/2+2+\varepsilon}(\mathbb{R}^{2n})$ . Expanding the function  $\psi_z((x+y)/2, \xi)$  by Taylor's formula at  $x = y$  and  $y = x$ , replacing  $(x-y)e^{ih^{-1}(x-y)\cdot\xi}$  with  $-ih \nabla_\xi e^{ih^{-1}(x-y)\cdot\xi}$  in (6) and integrating by parts with respect to  $\xi$ , one can easily prove that  $\mathcal{Q}_{1/2}(\varphi_z(\sigma)) - \mathcal{Q}_1(\psi_z(\sigma)) \mathcal{Q}_0(\psi_z(\sigma))$  coincides with a finite sum of operators  $R_k$  whose Schwartz kernels are given by oscillatory integrals of the form

$$i h^{1-n} \int_{\mathbb{R}^n} \int_0^1 \int_0^1 e^{ih^{-1}(x-y)\cdot\xi} a_{k,1}(l_{k,1}(t_1, x, y), \xi) a_{k,2}(l_{k,2}(t_2, x, y), \xi) dt_1 dt_2 d\xi,$$

where  $a_{k,j}$  is a derivative of  $\psi_z(\sigma)$  of order 0, 1 or 2 and  $l_{k,j}(t, x, y)$  is one of the following functions:  $x$ ,  $y$ ,  $x + t(y - x)/2$  or  $y + t(x - y)/2$ . The amplitudes in these oscillatory integrals vanish for all sufficiently large  $\xi$  because the functions  $\psi_z(\sigma)$  are constant outside  $D$  and each amplitude contains at least one derivative of  $\psi_z$  of order 1 or 2. Furthermore, we are only interested in  $x, y \in M$ . Therefore we can replace  $a_{k,j}(l_{k,j}(t_j, x, y), \xi)$  with  $b_{k,j}(l_{k,j}(t_j, x, y), \xi) := \chi(l_{k,j}(t_j, x, y), \xi) a_{k,j}(l_{k,j}(t_j, x, y), \xi)$  where  $\chi$  is a  $C_0^\infty$ -function on  $\mathbb{R}^{2n}$  which is equal to one on a sufficiently large ball. We have

$$\begin{aligned} & b_{k,1}(l_{k,1}(t_1, x, y), \xi) b_{k,2}(l_{k,2}(t_2, x, y), \xi) \\ &= (2\pi)^{-2n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{i\eta_1 \cdot l_{k,1}(t_1, x, y)} \hat{b}_{k,1}(\eta_1, \xi) e^{i\eta_2 \cdot l_{k,2}(t_2, x, y)} \hat{b}_{k,2}(\eta_2, \xi) d\eta_1 d\eta_2, \end{aligned}$$

where  $\hat{\cdot}$  denotes the Fourier transform with respect to the first  $n$  variables. Since  $b_{k,j} \in C_0^{n/2+\varepsilon}(\mathbb{R}^{2n})$ , it follows that the functions  $(1 + |\eta|)^{n/2+\varepsilon} \hat{b}_{k,j}(\eta, \xi)$  belong to  $L_2(\mathbb{R}^{2n})$  and their  $L_2$ -norms are estimated by constants depending on  $D$  and the Hölder norms of  $\sigma$  and  $\psi_z$ . Substituting this representation in

the corresponding oscillatory integral, we see that  $R_k = R_{k,1}^* R_{k,2}$  where the  $R_{k,j}$  are Hilbert-Schmidt operators acting from  $L_2(M)$  to  $L_2(\mathbb{R}^{3n} \times [0, 1]^2)$ .

Since  $\mathcal{Q}_1(\psi_0(\sigma)) \in \mathfrak{S}_2$  and  $R_{k,j} \in \mathfrak{S}_2$ , we have  $\mathcal{Q}_{1/2}(\varphi_0(\sigma)) \in \mathfrak{S}_1$ . The Hilbert-Schmidt norms of  $R_{k,j}$  are bounded by  $C_{k,j} h^{1-n}$  where  $C_{k,j}$  are constants depending only on  $\delta$ ,  $M$ ,  $D$ , and the Hölder norms of  $\sigma$  and  $\varphi$ . Therefore Theorem 7 implies that

$$\mathrm{Tr} \varphi_0(\mathcal{Q}_{1/2}(\sigma)) \leq (2\pi h)^{-n} \int_{T^*M} \varphi_0(\sigma) dx d\xi + C h^{1-n},$$

where  $C$  is a constant depending on the same parameters. If  $\mathcal{Q}_{1/2}(\sigma) \in \mathfrak{S}_1$  and  $\varphi(0) = 0$  then the same estimate holds for the function  $\varphi$ .

## REFERENCES

- [1] F. Berezin. *Convex functions of operators*. Mat. Sb. **88** (1972), 268–276 (Russian).
- [2] F. Berezin. *Covariant and contravariant symbols of operators*. Izv. Akad. Nauk SSSR Ser. Mat. **36** (1972), 1134–1167 (Russian).
- [3] M. Dimassi and J. Sjöstrand. *Spectral Asymptotics in the Semiclassical Limit*. Cambridge University Press, 1999.
- [4] W.D. Evans, R.T. Lewis, H. Siedentop and J.P. Solovej *Counting eigenvalues using coherent states with an application to Dirac and Schrödinger operators in the semiclassical limit*, Ark. Mat. **34** (1996), 265–283.
- [5] I. Ekeland and R. Temam. *Convex analysis and variational problems*. Studies in Mathematics and Its Applications, vol. 1. North-Holland Publishing Company and American Elsevier Publishing Company, 1976. Russian translation: “Mir”, 1979.
- [6] J.R. Giles. *Convex analysis with application in the differentiation of convex functions*. Research Notes in Mathematics, vol. 58. Pitman, 1982.
- [7] L. Hörmander. *The Analysis of Linear Partial Differential Operators III*. Springer-Verlag, 1985. Russian translation: “Mir”, 1987.
- [8] A. Laptev and Yu. Safarov. *A generalization of the Berezin–Lieb inequality for convex functions*, Amer. Math. Soc. Transl. Ser. 2 **175** (1996), 69–80.
- [9] A. Laptev and Yu. Safarov. *Szegő type limit theorems*, J. Funct. Anal. **138** (1996), 544–559.
- [10] Y. Safarov. *Pseudodifferential operators and linear connections*. Proc. London Math. Soc., **74** (1997), 379–417.
- [11] Yu. Safarov and D. Vassiliev. *The asymptotic distribution of eigenvalues of partial differential operators*. American Mathematical Society, 1996.
- [12] D. Tataru. *On the Fefferman–Phong inequality and related problems*. Comm. Partial Differential Equations **27** (2002), no. 11–12, 2101–2138.
- [13] N. Lerner and Y. Morimoto. *On the Fefferman–Phong inequality and a Wiener-type algebra of pseudodifferential operators*. Preprint 2005, <http://www.perso.univ-rennes1.fr/nicolas.lerner/FPI.2005.pdf>  
(The reference was added after the paper had been sent to print, and therefore does not appear in the published text.)

DEPARTMENT OF MATHEMATICS, KING’S COLLEGE, STRAND, LONDON WC2R 2LS  
E-mail address: ysafarov@mth.kcl.ac.uk